Structure
Article

Structural Basis for Substrate Targeting and Catalysis by Fungal Polysaccharide Monooxygenases

Xin Li,[1] William T. Beeson IV,[2] Christopher M. Phillips,[3] Michael A. Marletta,[2,3,4,5] and Jamie H.D. Cate[2,3,4,*]
[1]California Institute for Quantitative Biosciences
[2]Department of Chemistry
[3]Department of Molecular and Cell Biology
University of California, Berkeley, CA 94720, USA
[4]Division of Physical Biosciences, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[5]Present address: The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, CA 92037, USA
*Correspondence: jcate@lbl.gov
DOI 10.1016/j.str.2012.04.002

SUMMARY

The use of cellulases remains a major cost in the production of renewable fuels and chemicals from lignocellulosic biomass. Fungi secrete copper-dependent polysaccharide monooxygenases (PMOs) that oxidatively cleave crystalline cellulose and improve the effectiveness of cellulases. However, the means by which PMOs recognize and cleave their substrates in the plant cell wall remain unclear. Here, we present structures of *Neurospora crassa* PMO-2 and PMO-3 at 1.10 and 1.37 Å resolution, respectively. In the structures, dioxygen species are found in the active sites, consistent with the proposed cleavage mechanism. Structural and sequence comparisons between PMOs also reveal that the enzyme substrate-binding surfaces contain highly varied aromatic amino acid and glycosylation positions. The structures reported here provide evidence for a wide range of PMO substrate recognition patterns in the plant cell wall, including binding modes that traverse multiple glucan chains.

INTRODUCTION

The recalcitrance of crystalline cellulose to hydrolysis and the complex ultrastructure of the plant cell wall remain significant barriers to the efficient and economic conversion of lignocellulosic biomass to fermentable sugars (Galazka and Cate, 2011). Many microorganisms have evolved enzymes capable of degrading plant biomass to gain access to this abundant carbon source. Of these, a significant effort has been focused on filamentous fungi and the cellulases that they secrete. Cellulases are hydrolytic enzymes that cleave the β-1-4 glycosidic bonds present in cellulose. For hydrolysis to occur, glucan chains must be separated from the crystalline cellulose surface and threaded into a tunnel or cleft containing the cellulase active site. This chain separation process has been proposed to be a rate-limiting step for cellulose hydrolysis, as hydrolysis of amorphous forms of cellulose is significantly faster compared to crystalline substrates (Wilson, 2009).

For many years hydrolases were the only enzymes known to depolymerize cellulose, although many other proteins have been implicated in cellulose degradation. One such class of enzymes was referred to as glycosyl hydrolase family 61 enzymes (GH61s). These enzymes have been identified in the genome, transcriptome, and secretome of multiple cellulolytic fungi (Berka et al., 2011; Eastwood et al., 2011; Tian et al., 2009) and have been shown to significantly enhance the activity of cellulases (Harris et al., 2010; Langston et al., 2011; Phillips et al., 2011; Quinlan et al., 2011). The first structure of a GH61 enzyme revealed a flat substrate-binding surface lacking the conserved carboxylic acid residues essential for hydrolytic cleavage by cellulases (Karkehabadi et al., 2008).

Recently, it was shown that GH61 enzymes are copper-dependent enzymes that use reducing equivalents provided by the enzyme cellobiose dehydrogenase (CDH) or small molecule reductants to catalyze the oxidative cleavage of cellulose (Langston et al., 2011; Phillips et al., 2011; Quinlan et al., 2011). We proposed a mechanism of action for these enzymes resulting in oxygen insertion into cellulose at C-H bonds adjacent to the glycosidic linkage (Beeson et al., 2012; Phillips et al., 2011). The hydroxylated product is unstable and decomposes with the elimination of the adjacent carbohydrate moiety, thereby causing a break in the glucan chain. Hence, these enzymes are polysaccharide monoxygenases (PMOs). A larger family of proteins, such as the chitin- and cellulose-degrading CBM33 enzymes, have less than 10% sequence identity but may belong to the same superfamily (Forsberg et al., 2011; Vaaje-Kolstad et al., 2010). The fungal PMOs have been shown to yield cellulose cleavage products oxidized at either the C1 or C4 position to form an aldonolactone or 4-ketoaldose, respectively (Beeson et al., 2012; Phillips et al., 2011). Based on the sequence similarity and the product specificity, PMOs can be classified into at least three types (Figure S1A available online). Type-1 PMOs hydroxylate the C1 position of pyranose residues and produce an aldonolactone. Type-2 PMOs hydroxylate the C4 position of pyranose residues and produce a 4-ketoaldose. Type-3 PMOs appear to have less specificity, producing an aldonolactone and/or nonreducing end oxidized products (Beeson et al., 2012; Phillips et al., 2011; Quinlan et al., 2011).

Table 1. Data Collection and Refinement Statistics

	PMO-2	PMO-3
Data collection		
Space group	P21 (4)	P21 (4)
Cell dimensions		
a, *b*, *c* (Å)	67.17, 41.99, 69.25	35.74, 77.89, 82.05
α, β, γ (°)	90.00, 97.94, 90.00	90.00, 90.02, 90.00
Resolution (Å)	22.86–1.10 (1.16–1.10)[a]	35.74–1.37 (1.45–1.37)
R_{sym}	0.041 (0.409)	0.071 (0.475)
$I / \sigma I$	9.7 (1.7)	7.0 (1.6)
Completeness (%)	97.0 (94.9)	94.6 (93.5)
Redundancy	2.4 (2.5)	2.8 (2.8)
Refinement		
Resolution (Å)	22.86–1.10 (1.13–1.10)	35.74–1.37 (1.41–1.37)
No. reflections	142,387 (10,069)	84,275 (5,976)
R_{work} / R_{free}	0.132/0.149 (0.323/0.353)	0.116/0.149 (0.246/0.310)
No. atoms		
Protein	6,577	6,679
Ligand/ion	92	116
Water	538	759
B factors		
Protein	14.3	13.1
Ligand/ion	21.6	20.0
Water	29.3	27.2
Rmsds		
Bond lengths (Å)	0.0087	0.0088
Bond angles (°)	1.342	1.380
Ramachandran plot statistics		
Favored residues (%)	95.9	95.7
Allowed residues (%)	3.2	3.4
Outlier residues (%)	0.9	0.9

See also Figure S2.
[a]Values in parentheses are for highest-resolution shell.

Remarkably, the sequence and phylogenetic diversity of fungal PMOs is quite large. For example, the genomes of many cellulolytic fungi encode more PMOs than cellulases (Berka et al., 2011; Eastwood et al., 2011; Espagne et al., 2008; Ohm et al., 2010), suggesting that oxidative cleavage of cellulose is a major contributor to lignocellulose degradation. In *Neurospora crassa,* the ten PMO enzymes upregulated in response to growth on biomass (Tian et al., 2009) have an average pairwise sequence identity of only 33%. Here we present atomic-resolution structures of two PMOs from *N. crassa*, NCU01050 (PMO-2) and NCU07898 (PMO-3), which provide structural insight into the reaction mechanisms of PMOs. Structural and sequence comparisons also reveal the diversity of substrate binding sites presented by PMOs. These findings will aid in the identification of PMO substrates, and in the development of improved enzyme cocktails for deconstruction of the plant cell wall.

RESULTS

Diversity of the Planar Surface Harboring the Active Site

Structures of *N. crassa* PMO-2 and PMO-3 were determined at 1.10 and 1.37 Å resolution, respectively (Table 1). They are single domain proteins that adopt an immunoglobulin-like β sandwich (Figures 1A and 1B), as previously observed in structures of *Hypocrea jecorina* Cel61B (HjCel61B) (Karkehabadi et al., 2008), *Thielavia terrestris* GH61E (TtGH61E) (Harris et al., 2010), and *Thermoascus aurantiacus* GH61A (TaGH61A) (Quinlan et al., 2011). The strands in the β sandwich are connected by eight loops with two or three short α helix insertions. Whereas the loops on one end of the β sandwich are short and taper to a conical tip, the loops on the other end of the β sandwich adjacent to the protein active site form a relatively flat surface (Figures 1A and 1B) that has been proposed to bind crystalline cellulose (Harris et al., 2010; Karkehabadi et al., 2008). Loop L2 is the most diverse region in the PMO family. Loop L2 varies in length as well as secondary structure, containing one to three helices in all five known PMO structures (Figure 1C). Notably, the close proximity of these loops to the active site of PMOs is topologically similar to two other secreted proteins, immunoglobulin (Saul and Poljak, 1992) and fibronectin III (Leahy et al., 1996), suggesting that all three protein families use the same loop regions for molecular recognition (Figure S1B).

PMOs are secreted proteins containing a number of post-translational modifications with structural and/or functional importance, including multiple disulfide bonds (Figure 1C) and sites of N-glycosylation. In the natively purified PMOs crystallized here, a single N-glycosylation site was found on each PMO, but at different locations. Whereas the N-glycosylation site in PMO-2 is distal from the predicted cellulose-binding surface, the glycan in PMO-3 extends onto the plane of the putative crystalline cellulose-binding surface (Figure 1B). In PMO-3, an N-acetylglucosamine dimer attached to Asn6 near the center of β strand S1 is clearly seen in electron density maps, with weak density extending from this dimer likely due to additional carbohydrate moieties. Mass spectra of PMO-3 from dissolved crystals and uncrystallized samples are identical and reveal that PMO-3 is heterogeneously glycosylated with five to ten monosaccharide units (Figure S1C). Thus, the glycan is long enough to form a significant extension of the planar surface and may play a role in substrate binding as suggested by recent molecular dynamic simulations performed on fungal cellulose binding module CBM1 (Taylor et al., 2012).

In the *N. crassa* PMO-3, the glycan on one side and a short 3-10 helix on the other side (H2.1, Figures 1B and 1C) form a shallow hydrophilic groove that is tilted toward the copper center (Figure 1B). By contrast, *N. crassa* PMO-2 has a shorter L2 loop, lacks the H2.1 helix, and its N-glycosylation sites are located on the other side of the molecule (Figures 1A and 1C). Notably, the H2.1 helix and the N-glycan that extends to the planar surface are either concurrent or both absent in all PMOs with known structure. PMO HjCel61B (Karkehabadi et al., 2008) has the same N-glycosylation site as *N. crassa* PMO-3, and contains a three-turn α helix at the position of H2.1. TaGH61A (Quinlan et al., 2011) has a slightly different arrangement and harbors the N-glycosylation site on another planar

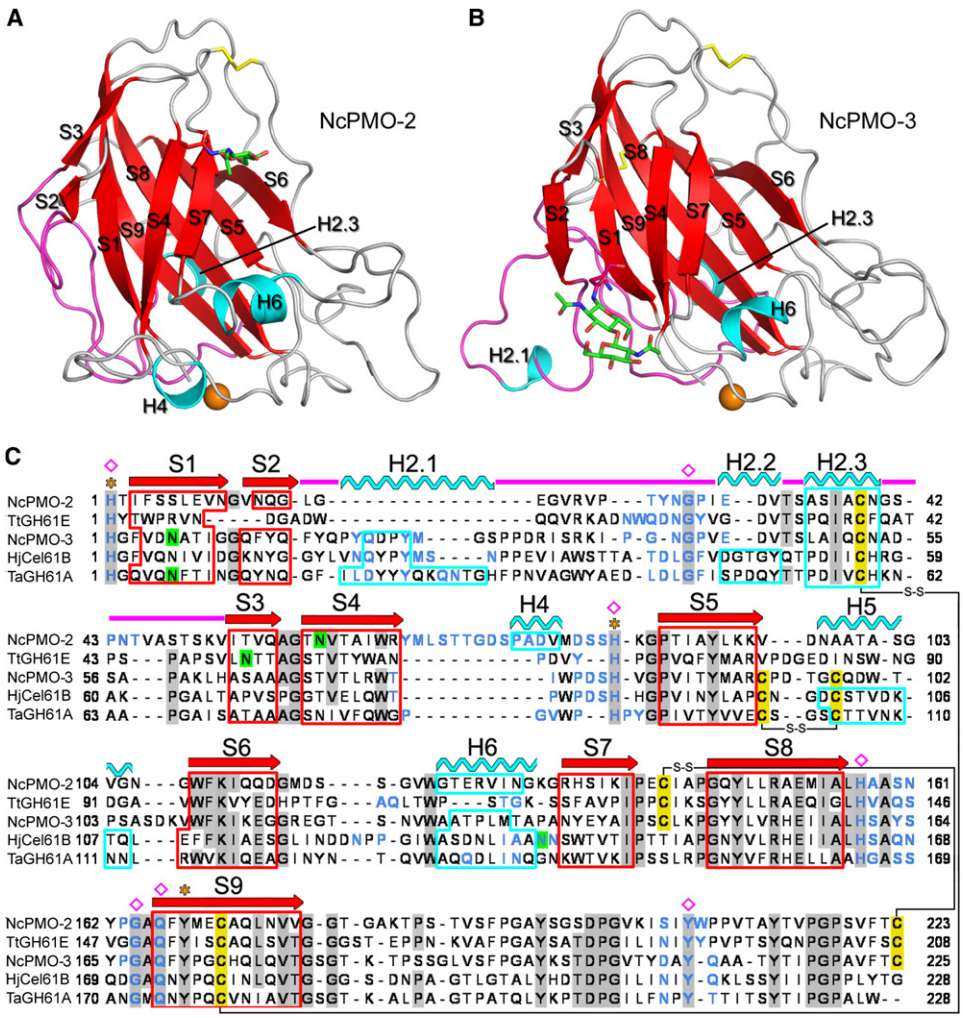

Figure 1. Structural Overview of PMO Family Enzymes

(A) A cartoon presentation of *N. crassa* PMO-2 with the planar surface facing down and the conical tip pointing up. Copper ion is shown as a brown sphere. Disulfides are in yellow. The glycosylated asparagine residue and glycan are shown as sticks with the carbon atoms in magenta and green, respectively. The highly variable loop L2 is colored magenta.

(B) PMO-3 from *N. crassa* viewed as in (A).

(C) Structure-based sequence alignment of *N. crassa* PMOs with other known X-ray crystal structures. Conserved residues are shaded in gray and the residues on the planar surface are colored in light blue. Conserved surface aromatic residues are labeled with pink diamonds. Segments assigned to β strands and α helices are framed with red and cyan lines, respectively, and are labeled in sequential order. Loop L2 is colored magenta. Cysteine residues are highlighted in yellow and linked with black lines to show disulfide bonding. Glycosylated asparagine residues are highlighted in green. Orange asterisks mark the copper-coordinating residues.

See also Figure S1.

surface residue (Asn138) and the H2.1 3-10 helix is conserved (Figure 1C). In contrast to HjCel61B and TaGH61A, TtGH61E (Harris et al., 2010) is similar to PMO-2, lacking the L2 loop and harboring its N-glycosylation sites distal to the planar surface (Figures 1A and 1C).

Copper Coordination and the Active Site

The active sites of the *N. crassa* PMOs reside at the center of the planar surface thought to bind to the substrate and coordinate a copper ion essential for catalysis (Phillips et al., 2011). As observed in TaGH61A (Quinlan et al., 2011), the essential N-terminal histidine is Nε-methylated, although the functional significance of this methylation remains unclear. The copper ion lies in an octahedral environment exhibiting Jahn-Teller distortion (Table 2; Jahn and Teller, 1937). The protein N-terminal amine, Nδ atom of Nε-Me-His1, Nε atom of His84 in PMO-2 (His82 in PMO-3), and a water occupy four corners of the equatorial plane which is tilted about 30° relative to the planar enzyme surface (Figure 2A). The three nitrogens coordinate the copper with typical bond distances, whereas the distance of the water oxygen from the copper varies widely. One of the axial positions is occupied by a buried tyrosine, Tyr168 in PMO-2 and Tyr171 in PMO-3, with the phenolic oxygen located 2.7–2.8 Å away from the copper (Table 2).

Table 2. Geometry of PMO Active Sites

	PMO-2/ Chain A	PMO-2/ Chain B	PMO-3/ Chain A	PMO-3/ Chain B
Cu-Nδ(HIC) (Å)[a]	1.92	1.92	1.95	1.90
Cu-N(HIC) (Å)	2.25	2.24	2.27	2.29
Cu-Nε(His) (Å)	1.99	1.97	2.11	2.09
Cu-water (Å)	1.84	1.82	3.60	2.31
Cu-O(Tyr) (Å)	2.76	2.80	2.73	2.62
Cu-O1 (Å)	2.96	2.92	3.44	N/A
O1-O2 (Å)[b]	1.16	1.15	1.49	N/A
Cu-O1-O2 (°)	147.75	142.11	117.16	N/A

See also Figure S6 and Table S1.

[a]The coordinate error for PMO-2 is ± 0.011 Å and for PMO-3 is ± 0.027 Å as determined by the standard uncertainties of the positional parameters based on the likelihood function (Murshudov et al., 1997).

[b]The O-O bond length was not restrained during the refinements in Table 1. The ideal O-O bond lengths in oxygen, superoxide, and peroxide are 1.21, 1.33, and 1.49 Å, respectively (Foote et al., 1995).

In the four unique PMO structures determined here (two per crystal form), the other axial coordination site is occupied by different ligands, providing structural evidence supporting the proposed mechanism used by PMOs in the oxidation of substrates. In the structure of PMO-2, the exposed axial position is occupied by an oblong-shaped electron density with an end-on (η1) configuration to the copper ion (Figure 2B). The electron density is best modeled as a dioxygen species (Figure S2A) with an unrestrained bond length of 1.16 Å. The bond length for oxygen or superoxide is 1.2–1.3 Å and peroxide would be close to 1.49 Å. This dioxygen species is consistent with a superoxide species weakly coordinated to Cu(II) (Table 2; Table S1), resulting from internal electron transfer in a previously formed Cu(I)-oxygen species as predicted as part of the PMO reaction mechanism (Klinman, 2006; Phillips et al., 2011; Solomon et al., 2011). Notably, the occupancy refinement shows that the proximal O1 atom has a higher occupancy (0.80, versus 0.71 of the O2 atom), which is consistent with the unequal electron distribution expected in a superoxide anion (Table S1). The projection of the O1-O2 bond on the equatorial plane of the copper coordination points to the N-terminal amine coordination in both chains, even though the crystal packing results in different microenvironments for the two active sites.

In the structures of PMO-3, the exposed axial position is either vacant (chain B, Figure 2D), or occupied by a dumbbell-shaped electron density 3.3 Å from the copper (chain A, Figure 2C). The shape of the electron density suggests a diatomic species, which is further supported by comparative modeling (Figure S2B). The distance between the centers of the dumbbell lobes of the electron density is 1.49 Å, a bond length expected for peroxide instead of molecular oxygen or superoxide (see Table 2). Given the long distance between the O1 and copper ion (Table 2), the peroxide species is most likely hydrogen

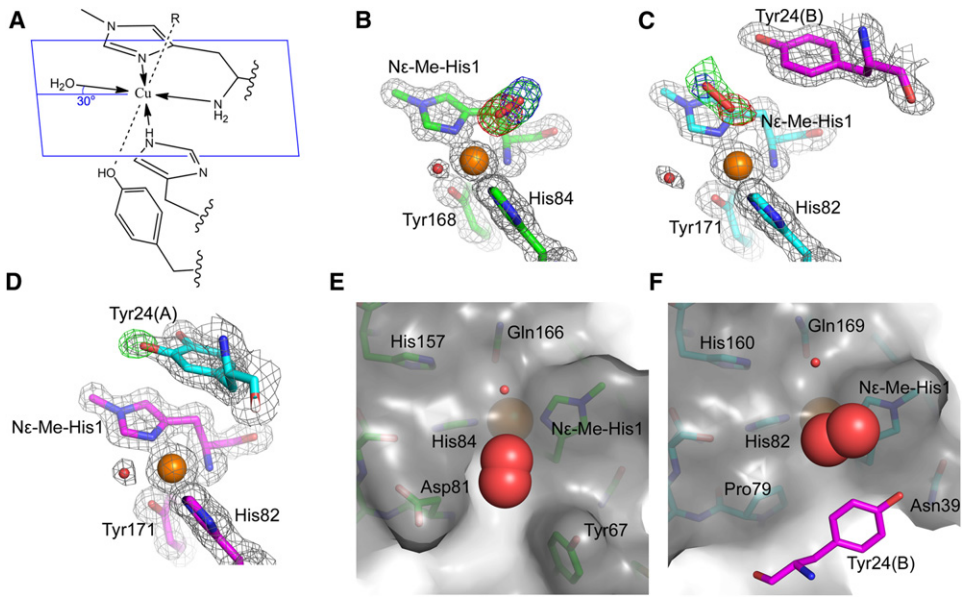

Figure 2. The Active Site of *N. crassa* PMOs

(A) A generalized schematic representation of the active site copper coordination. The blue lines denote the substrate binding plane. "R" is most consistent with superoxide in PMO-2 and with hydrogen peroxide in PMO-3 chain A.

(B–D) Weighted electron density maps surrounding the active site of PMO-2 chain A (B), PMO-3 chain A (C), and PMO-3 chain B (D) with $2mF_O$-DF_C (in gray, contoured at 1.2 σ). The mF_{Obs}-DF_{Calc} difference electron density maps after omitting the whole superoxide or hydrogen peroxide molecule are colored in green; after omitting only the proximal O1 atom are colored in red; and after omitting only the distal O2 atom are colored in blue. The Omit map of 3-hydroxyl group on Tyr24 from PMO-3 chain A is also colored in green and all omit maps are contoured at 3.0 σ. The orientation is approximately equivalent to that shown in the schematic in (A).

(E and F) Oxygen-binding grooves in PMO-2 (E) and PMO-3 (F). The positions of bound superoxide and peroxide are shown as red spheres.

See also Figure S3.

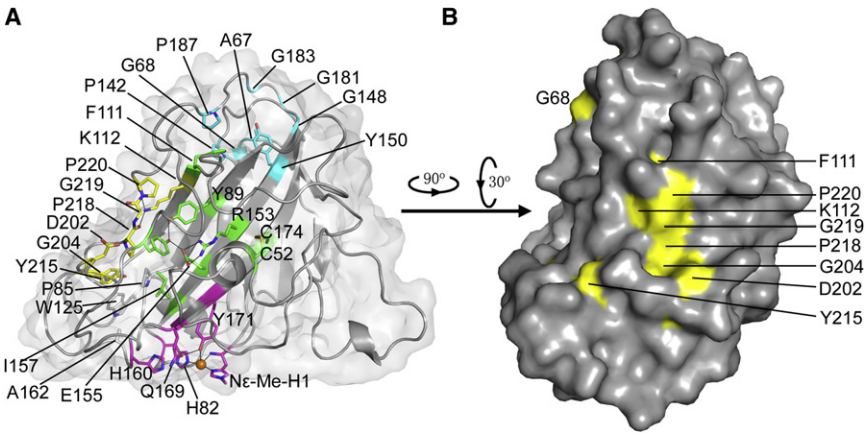

Figure 3. Conserved Residues across the PMO Family Mapped onto the Structure of *N. crassa* PMO-3

(A) PMO-3 is shown as cartoon presentation. The protein molecule is related to Figure 1B by rotating about 180 degrees around the vertical axis. The conserved residues with sequence identity higher than 85% are shown as sticks. Most of them can be grouped into four clusters located at the active site (in magenta), the conical tip (in cyan), a patch at the center of the hairpin-shaped C-terminal coil (in yellow), or on the larger sheet of the β sandwich fold (in green).

(B) Conserved residues on the surface of PMO-3 that connect to the putative electron transfer pathways are shown on the molecular surface in yellow, mapped onto the structure of *N. crassa* PMO-3.

See also Figure S4.

peroxide, held in place by additional interactions with the protein. The O1 atom of the modeled hydrogen peroxide is 3.09 Å from Nδ of Nε-Me-His1, and is hydrogen bonded to Ser81 through two waters. The OH group of Ser81 is 2.68 Å from Nδ of His82, another copper ligand. The O2 atom of the modeled hydrogen peroxide is also hydrogen bonded to a water molecule. The projection of the O-O bond on the equatorial plane of the copper coordination differs from that observed in the PMO-2 structures, and points to the methyl group attached to Nε-Me-His1.

To better understand why different dioxygen species are present in these PMO crystal structures, the microenvironments surrounding the active sites were compared. In PMO-2, the dioxygen species lies in a groove formed by side chains of Nε-Me-His1 and His84 (Figure 2E), and the O1-O2 bond runs almost parallel to the planar surface. The O1 atom is weakly coordinated with the copper ion that is positioned at the end of the groove and between the two copper-coordinating histidines. Residues Tyr67 and Asp81 extend the edges of the groove and further restrain the radial orientation of the O2 atom. Whereas the two histidines and the copper are conserved in all PMOs, Asp81 in PMO-2 is replaced with Pro79 in PMO-3, and Tyr67 in PMO-2 is missing in PMO-3. Thus, PMO-3 has a shallower groove on its surface compared to PMO-2 (Figure 2F).

In the asymmetric unit of crystals of PMO-3, two enzymes are packed against each other by means of their planar surfaces via a pseudo 2-fold noncrystallographic axis. The contact area is dominated by Nε-Me-His1, Tyr20, and Tyr24 from each monomer. The aromatic rings of all three residues are almost parallel to the planar surface, stacking on the same triplet from the other monomer (Figure S3A). This packing pattern resembles how cellulose-binding modules, like CBM1, are thought to use three conserved aromatic residues to bind three glucose units on the crystalline cellulose surface (Boraston et al., 2004; Kraulis et al., 1989; Figure S3B). As a result of this packing, Tyr24 of each monomer is positioned close to the active site of the neighboring monomer. We suggest that the phenol ring of Tyr24 on the planar surface of PMO-3 mimics the pyranose ring of a glucose unit that would be present in the PMO substrate on the surface of crystalline cellulose. Further evidence suggesting that Tyr24 mimics the substrate comes from the observation that Tyr24 of chain A is oxidized to 3,4-dihydroxyphenylalanine by the active site of chain B after crystal exposure to X-rays (Figure 2D). The conversion is unlikely to be caused directly by X-ray radiation, because all the other 15 tyrosines are unmodified and radiation damage removes the hydroxyl group (Burmeister, 2000; Leiros et al., 2001, 2006) or decreases Cζ density (Leiros et al., 2001). The hydroxylation of Tyr24 in one of the monomers likely occurred due to electron capture in the adjacent enzyme active site during X-ray exposure that served to activate a bound dioxygen species for the hydroxylation reaction. Consistent with this hypothesis, mass spectra revealed a 16 Da increase in PMO-3 mass in samples extracted from crystals exposed to X-rays, but not in samples from crystals prior to X-ray exposure, or in samples from crystals of copper-depleted PMO-3 exposed to X-rays (Figure S3C). Tyr24 hydroxylation is consistent with the monooxygenase activity of PMOs although the chemical mechanism for tyrosine oxidation may differ from that in glucan chains due to the different C-H bond dissociation energies for phenols and carbohydrates (Blanksby and Ellison, 2003).

Mapping Globally Conserved Residues in the PMO Family

The conserved copper-containing active site suggests that a similar reaction mechanism is used by all members in the fungal PMO family. To identify other residues important for the reaction, the sequences of other cellulose-induced PMOs from *N. crassa* were used in a structure-based sequence alignment (Figure S4A) and conserved residues were mapped onto the structure of PMO-3 to examine their locations and roles (Figure 3A). Approximately 15% of residues are highly conserved with greater than 85% sequence identity. These residues can be grouped into four clusters located at the active site, the conical tip, a patch at the center of the hairpin-shaped C-terminal coil, or on the larger sheet of the β sandwich fold. While the cluster at the conical tip likely serves a structural role, the other three may participate directly in substrate recognition and the electron transfer required for the oxidation reaction.

On the putative substrate-binding plane, only the residues surrounding the copper are conserved, including the copper ligands (Nε-Me-His1/His82/Tyr171, PMO-3 numbering) and the two adjacent residues Gln169 and His160. The second cluster

is an internal ionic network centered by a buried salt bridge between Arg153 and Glu155 that connects the copper ligand Tyr171 with the surface-exposed Lys112 via conserved Tyr198 and Tyr89 (PMO-3 numbering). Two aromatic residues, Trp126 and Tyr215, adjacent to His160 form the third cluster. Interestingly, the last two clusters both connect the active site to a conserved surface patch on the edge of the β sandwich. The patch is fairly hydrophilic with a lysine and an aspartic acid separated by a Pro-Gly-Pro motif (Figure 3B) providing a possible recognition site for binding a partner protein such as CDH.

Subfamily-Conserved Residues

Apart from the few amino acids described above, the various PMOs secreted by each filamentous fungus during growth on cellulose are highly sequence divergent. For instance, of the 10 *N. crassa* PMOs upregulated during growth on cellulose (Tian et al., 2009), the pairwise sequence identities range from 21% to 64%. Sequence alignments of *N. crassa* PMOs to those from other fungi with known structures range from 24% to 39%. Different PMOs have been shown to generate different products corresponding to the oxidation of the C1 or C4 carbon of the pyranose ring in cellulose chains (Beeson et al., 2012; Phillips et al., 2011). One possible explanation is that PMOs bind to cellulose uniquely such that the active site is in proximity to different carbons for hydrogen atom abstraction. We performed multiple sequence alignments using members of the PMO-2 and PMO-3 subfamilies within type-2 and type-3 PMOs, respectively (Figures S1A and S5), and mapped the sequences to the respective structures (Figures 4A and 4B). In the subfamily represented by PMO-2, 21% of the amino acid positions are occupied by identical residues. In the PMO-3 subfamily, 19% of the residues are occupied by identical amino acids. The conserved residues cover most of the substrate-binding plane, with the globally conserved residues positioned at the center of the plane and the subfamily-specific residues scattered around it. Notably, the subfamily-specific residue distributions display distinct patterns in the PMO-2 and PMO-3 subfamilies (Figures 4A and 4B).

Fungal cellulases often include a cellulose-binding module (CBM1), which has a flat surface containing aromatic residues important for crystalline cellulose binding (Boraston et al., 2004; Kraulis et al., 1989). In the NMR structure of CBM1 from *H. jecorina* Cel7A, three tyrosines are arranged in a line on a planar surface and are separated from each other with a distance of 10.5 or 10.6 Å corresponding to the length of two β-1-4 linked glucose units in a single glucan chain (Figure 4C; Kraulis et al., 1989). By contrast, four conserved aromatic residues Nε-Me-His1, Tyr20, Tyr24, and Tyr210 span the flat surface of *N. crassa* PMO-3 and have a different spatial distribution compared to those on CBM1. The distance between Tyr20 and Tyr24 is 8.3 Å, which matches closely to the interchain distance of cellulose (8.2 Å) (Nishiyama et al., 2002). When placing these two residues on the pyranose rings from two neighboring glucan chains, Nε-Me-His1 and Tyr210 would also stack with other pyranose rings (Figure 4C). Thus, the distribution pattern of the surface aromatic residues supports an interchain binding mode for the PMO-3 subfamily. Importantly, while Nε-Me-His1 and Tyr210 are conserved in nearly all PMOs, Tyr20 and Tyr24 are conserved in every PMO-3 but not in any member of the PMO-2 subfamily (Figure S5).

In contrast to PMO-3 subfamily, five aromatic residues are present on the planar surface of the PMO-2 subfamily (Figure 4A; Figure S5). Only two of them, Nε-Me-His1 and Tyr206, have their aromatic rings relatively parallel to the planar surface of the enzyme, which prevents accurate docking of the remaining aromatic residues to the planar surface of the cellulose lattice. Nevertheless, the different location of these aromatic residues suggests that the PMO-2 subfamily binds cellulose in a different orientation compared to the PMO-3 subfamily. Notably TtGH61E, a close structural homolog of *N. crassa* PMO-2 from a type-1 PMO subfamily (Figure S1A), has three tyrosines arranged in a line on the planar surface of the enzyme with distances of 10.4 and 14.2 Å between adjacent Tyr residues (Harris et al., 2010). These three Tyr residues can be stacked on three pyranose residues (n, n+2, n+5) in a single glucan chain in crystalline cellulose, consequently positioning His1 adjacent to the neighboring glucan chain (Figure 4C). In the case of TaGH61A (Quinlan et al., 2011), a member of the type-3 PMOs (Figure S1A), two surface tyrosines aligned with His1 span distances of 15.5 and 16.5 Å from His1, respectively. These three residues are also arranged in a line and can be stacked on three pyranose residues in a single glucan chain in a (n, n+3, n+6) manner. In this case, the active site is positioned near the same glucan chain (Figure 4C).

DISCUSSION

All fungal PMO enzymes may use the same oxidative chemistry to overcome the recalcitrance of the crystalline cellulose substrate as evidenced by the identical copper binding sites in five fungal PMO structures. The dioxygen species and hydroxylation product observed in the present structures of *N. crassa* PMO-2 and PMO-3, likely due to in situ reduction of the active sites by X-ray exposure, provide additional support for the oxygenation reaction mechanism proposed previously (Phillips et al., 2011) and recently further substantiated (Beeson et al., 2012). In the first step of the proposed mechanism (Figure S6), copper reduction and oxygen binding are followed by internal electron transfer to form a PMO-Cu^{II} superoxo complex that is consistent with the modeled superoxide species seen in the *N. crassa* PMO-2 structure (Figure 2B). In the next step, the copper superoxo may directly abstract a hydrogen atom from the nearby substrate, generating a copper hydroperoxo intermediate and a substrate radical. In the third step of the reaction, a second electron could then facilitate O-O bond cleavage in the copper peroxo species, releasing water and generating a copper oxo radical that could couple with the substrate radical to hydroxylate the polysaccharide. Finally, hydroxylation at the C1 or C4 position of pyranose moieties will destabilize the glycosidic bond (Beeson et al., 2012; Phillips et al., 2011) with the elimination reaction possibly catalyzed by a conserved third histidine present in all PMOs (His 157 in PMO-2 and His160 in PMO-3, Figures 3A and 4). Cleavage of the glycosidic bond by PMOs would thereby generate a regular glycan and an aldonolactone or 4-ketoaldose. The structure of PMO-3 chain A with a modeled hydrogen peroxide in the active site may represent an intermediate state of the reaction cycle, and the hydroxylation of Tyr24 in chain B may be an outcome of oxygen reduction (Figure 2C). Tyr24 hydroxylation may follow a different reaction mechanism

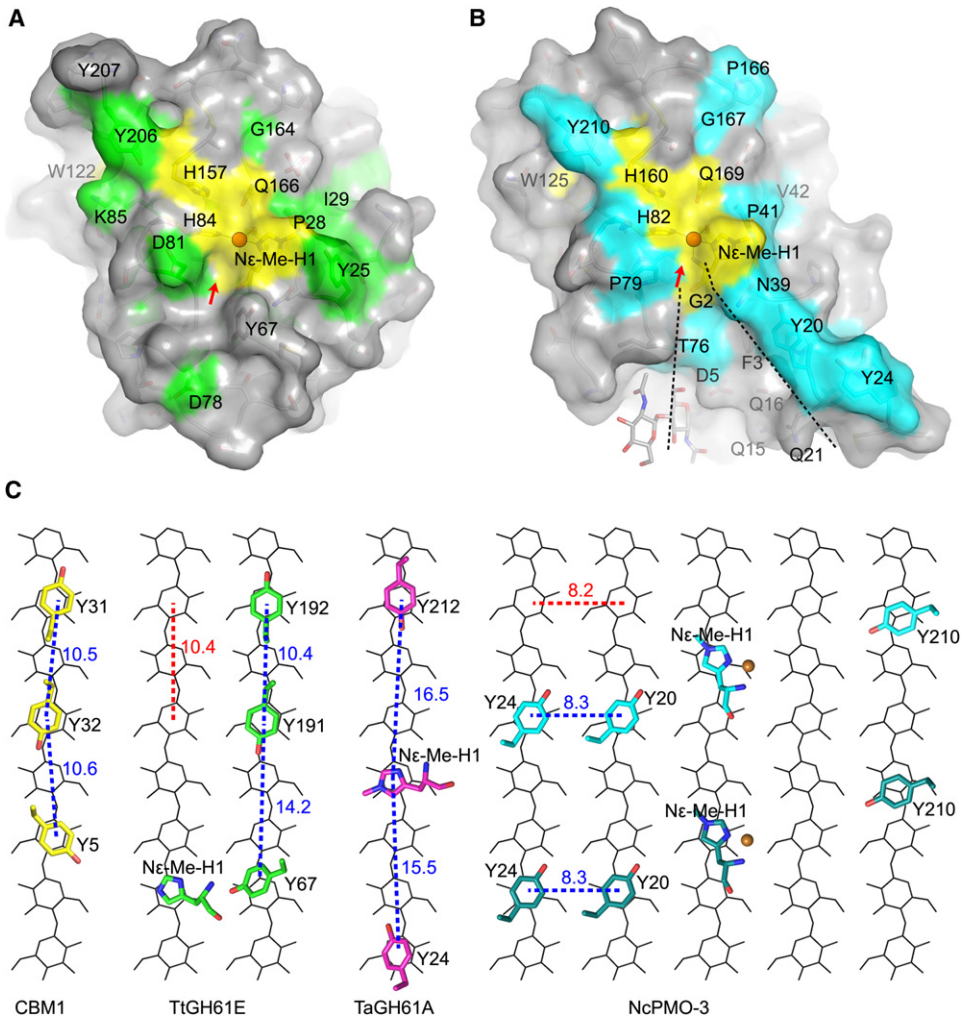

Figure 4. Models for Substrate Binding by PMOs

(A) Planar surface of PMO-2 showing conserved residues. The globally conserved and the PMO-2 subfamily conserved residues are colored in yellow and green, respectively. The copper ion is shown as an orange sphere. The red arrow indicates the oxygen binding groove.

(B) Planar surface of PMO-3 showing conserved residues in PMO-3, as described for PMO-2 in (A), with PMO-3 subfamily residues colored cyan. Dashed lines outline the shallow hydrophilic groove between the H2.1 and the N-glycan in *N. crassa* PMO-3.

(C) Docking of the flat surface aromatic residues of a CBM1 and PMOs to the crystalline cellulose hydrophobic face built from the crystal structure of cellulose Iβ (Nishiyama et al., 2002). Cellulose chains are shown as black lines and the distances between pyranose units are labeled with red dotted lines. Aromatic residues on the flat surface of each protein are shown as sticks and the distances are labeled with blue dotted lines. The carbon atoms of TrCel7A-CBM1 (PDB entry 1CBH), TtGH61E (PDB entry 3EJA), TaGH61A (PDB entry 3ZUD), and NcPMO-3 are colored in yellow, green, magenta, and cyan, respectively. Alternative binding of each protein can be formed by moving along the glucan chain for a distance of an odd number of pyranose units, as shown in the figure for two PMO-3 s, in dark and light cyan, respectively, and separated from each other by three pyranose units along the glucan chain.

See also Figure S5.

than the one proposed above, but it demonstrates the generation of a reactive oxygen species when electrons are supplied. These observed structural features, including the modeled superoxide in PMO-2 and the modeled hydrogen peroxide and hydroxylated Tyr24 in PMO-3, are indicative of the types of chemistry that can be supported by PMOs and are consistent with the proposed catalytic mechanism.

In addition to forming a highly reactive copper center for oxidative polysaccharide degradation, PMOs face the challenge of accessing their substrates in the complex structure of the plant cell wall. The cell wall contains other components like lignin, hemicellulose, and pectin. Cellulose itself displays structural polymorphism including different crystalline forms (Larsson et al., 1997) and twisted microfibrils (Fernandes et al., 2011; Matthews et al., 2006), which will lead to highly diversified cellulose surfaces exposed on the plant cell wall. To degrade such a diversity of targets, filamentous fungi like *N. crassa* secrete a large number of PMOs with different structural features on their predicted substrate recognition surfaces. First, the aromatic residues on the planar surface are highly divergent in number and location, as exemplified in *N. crassa* in which none of these residues is absolutely conserved in its ten PMOs upregulated by

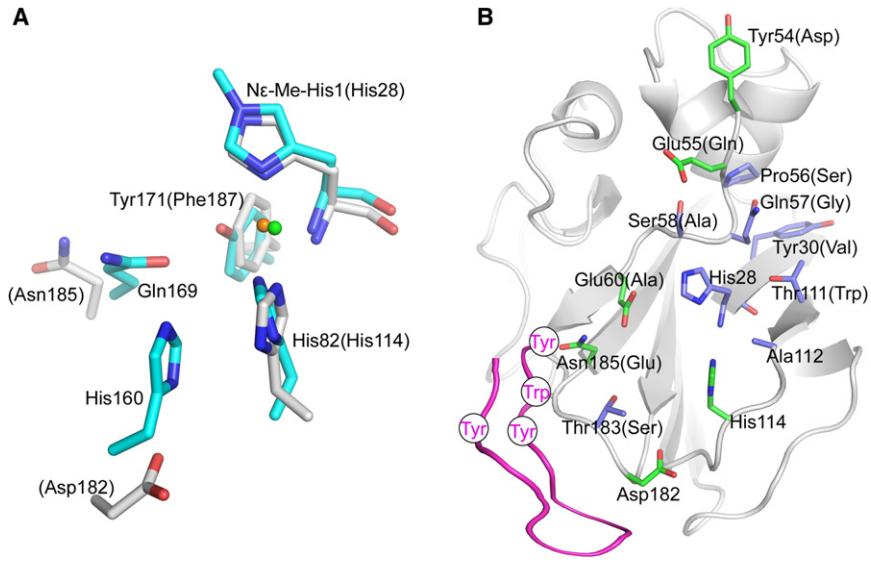

Figure 5. Active Site and the Substrate Binding Surface in CBM33 Family

(A) Active sites alignment of *N. crassa* PMO-3 (in cyan) and *S. marcescens* CBP21 (PDB entry 2BEM, in gray, residue numbers in parentheses). While His1 and His82 are conserved in CBP21, the axial coordinating residue tyrosine, which is absolutely conserved in all PMOs, is replaced with a phenylalanine in many CBM33s. Furthermore, Gln169 and His160 (PMO-3 numbering) are conserved in all PMOs but are not conserved in CBP21. His160 may catalyze the final elimination reaction to cleave the glycosidic bond and is replaced with an aspartic acid in CBP21 and many other CBM33 proteins.

(B) Substrate binding surface in the CBM33 family. The *S. marcescens* CBP21 structure (PDB entry 2BEM) is shown in gray cartoon, with the conserved residues on the flat surface shown as sticks. The conserved hydrophilic residues that have been experimentally proven to be important for chitin binding (Vaaje-Kolstad et al., 2005) are colored in green. In *S. coelicolor* CelS2, which acts on cellulose (Forsberg et al., 2011), differences in residues arrayed on the planar surface are labeled in parentheses. A sequence insertion in CelS2 that contains four aromatic residues (circles), is modeled schematically in magenta.

See also Figure S7.

biomass. Second, the PMO N-glycosylation sites are located at different positions, with some of the N-glycans likely serving as an integral component of the substrate binding planes, as seen in *N. crassa* PMO-3. Third, while all reported structures of PMOs lack a dedicated cellulose-binding domain, C-terminal CBM1 domains are frequently found in fungal PMOs that would play a role in substrate recognition and binding. Finally, PMOs may have binding grooves on their surface for molecular oxygen, as in PMO-2, to promote specific hydrogen atom abstraction adjacent to the glycosidic bond.

Once PMOs bind to the cellulose surface, their active sites, located at the center of the substrate-binding plane, are probably inaccessible to external electron donors and may require long distance electron transfer during the reaction cycle. Hydrogen bonds and π orbitals are the most likely electron tunneling media in proteins (Gray and Winkler, 1996), and both may play a role in electron transfer in PMOs. Fungal PMOs have an internal conserved ionic network centered with a buried salt bridge between Arg153 and Glu155 that connects copper-ligand Tyr171 with the surface-exposed Lys112 via conserved Tyr198 and Tyr89 (PMO-3 numbering). Long distance electron transfer may be carried through this conserved hydrogen bond network. Another electron path may be provided by two nearby aromatic residues, Trp126 and Tyr215, adjacent to His160. Both potential pathways are composed of conserved residues across all PMOs and connect the active site to a relatively hydrophilic patch on the surface of the protein (Figure 3B). CDH, a fungal secreted flavocytochrome, is a known electron donor to PMOs (Langston et al., 2011; Phillips et al., 2011). Notably, the heme-containing domain of CDH from a distantly related basidiomycete fungus (Hallberg et al., 2000) can be docked computationally to the conserved hydrophilic surface patch (Figure S4B). Thus, the electrons required to reduce copper-oxygen species could be transferred through these two conserved paths in fugal PMOs from the CDH heme-containing domain to the PMO active site.

Proteins with structural similarity to the fungal PMOs can be found in bacteria as CBM33 family proteins defined in the CAZy database (Cantarel et al., 2009) although these proteins generally have less than 10% sequence identity when compared to PMOs in *N. crassa*. The best characterized CBM33 member, *Serratia marcescens* CBP21, has been shown to degrade crystalline chitin in the presence of divalent metal ions and oxygen (Vaaje-Kolstad et al., 2010). Chitin and cellulose are both recalcitrant polysaccharides and their pyranose backbone chains are similar. However, chitin has an acetamido group replacing a hydroxyl group in cellulose on the ring C2 position, which renders the surface of chitin more hydrophilic and amenable to polar interactions. Comparison between these two families provides further hints regarding substrate recognition and the reaction mechanism. The protein database (Protein Data Bank [PDB]) contains crystal structures of three CBM33 members, including *S. marcescens* CBP21 (PDB entry 2BEM) (Vaaje-Kolstad et al., 2005) and two other CBM33s (from *Burkholderia pseudomallei* and *Vibrio cholera*, PDB entries 3UAM and 2XWX respectively, with no associated publication). CBM33s also contain a central β sandwich fold but have more helical secondary structures than PMOs. Notably, of five absolutely conserved residues surrounding the active-site copper, only two histidines are retained in CBM33s (Figure 5A), which imply CBM33s may utilize a different molecular mechanism for polysaccharide oxidation. Whereas PMOs have multiple aromatic residues on the planar surface, CBM33s have many more hydrophilic residues on their flat surface (Figure 5B), and a hydrophilic patch in CBP21 has

been proposed for chitin recognition (Vaaje-Kolstad et al., 2005). The substrate binding surface properties of PMOs that target cellulose and CBM33s that target chitin therefore coincide with the different surface properties of cellulose and chitin.

Finally, it is worth noting that PMOs and CBM33s are highly diversified, and the substrate specificity to cellulose and chitin, respectively, is only based on a small number of tested enzymes. The *N. crassa* PMO-3 structure with the hydroxyl modification on Tyr24 suggests that the PMO superfamily may include enzymes with substrates that include phenol derivatives, such as those in lignin. Furthermore, the CBM33 member CelS2 from *Streptomyces coelicolor* has recently been shown to have cellulose oxidation activity (Forsberg et al., 2011). When compared to the three CBM33s with known structure, CelS2 has a cellulose-binding domain and a sequence insertion containing four aromatic residues (Figure 5B). These structural features may contribute to its binding to cellulose and cellulolytic activity. Future experiments to identify the substrates of the different PMO and CBM33 family members should reveal the wide array of polysaccharides that are targeted by this new class of monooxygenase enzymes, and should aid in the development of more efficient enzyme cocktails for low cost lignocellulosic biomass conversion.

EXPERIMENTAL PROCEDURES

Structure Determination

N. crassa PMO-2 and PMO-3 were purified from the secretome of *N. crassa* culture growing on cellulose (Avicel) as described previously (Phillips et al., 2011). PMO-2 crystals were grown at room temperature with polyethylene glycol 3350 as precipitant using the hanging-drop vapor diffusion method. The best PMO-3 crystals were obtained from a microcentrifuge tube incubated at 4°C for a week containing 50 μl of 10 mg/ml PMO-3 in 10 mM Tris-HCl, 100 mM NaCl (pH 8.5). X-ray diffraction data were collected at beamline 8.3.1 at the Advanced Light Source at Lawrence Berkeley National Laboratory. The initial phases were obtained by the molecular replacement method using search models built from the structures of HjCel61B (PDB entry 2VTC) (Karkehabadi et al., 2008) and TtGH61E (PBD entry 3EJA) (Harris et al., 2010). The models for PMO-2 and PMO-3 were built using the program COOT (Emsley and Cowtan, 2004) and were refined using simulated annealing with Phenix (Adams et al., 2002) to correct the model ambiguity in loop regions. The final model was refined with anisotropic temperature factors and all hydrogen atoms added using the program Refmac5 (Murshudov et al., 1997) within the CCP4 package using the CCP4i interface (Collaborative Computational Project, Number 4, 1994). PMO-3 dataset was twinned and was refined in Refmac5 with the twin refinement setting on, which allows automatic twin operator and twin fraction determination (twin operator: "-H, -K, L," fraction: 0.405). Table 1 summarized the statistics of data collection and refinement.

The O-O bond lengths of the dioxygen species in PMO-2 and PMO-3 were refined without restraints on bond length. In the PMO-2 refinement, the occupancies of the dioxygen species were first refined as a group, which would be consist with the electron distribution in molecular oxygen. However, in F_{obs}–F_{calc} difference electron density maps, negative density at a −3.5 σ level was found near the distal oxygen atom in both monomers in the asymmetric unit and the B factor of the copper-proximal oxygen atom O1 refined to a value higher than that of the distal oxygen O2. Therefore, the occupancies of the two oxygen atoms were refined individually. The final occupancies of O1 and O2 are 0.80 and 0.71, which is more consistent with the electron distribution in superoxide anion. In addition, the negative difference density near O2 was eliminated at a −3.0 σ level. Furthermore, the B factor of the copper-proximal oxygen atom O1 refined to a lower value than that of the distal oxygen O2 (Table S1).

Sequence Alignment of PMOs

The structure-based sequence alignment of *N. crassa* PMO-2, PMO-3, HjCel61B (PDB entry 2VTC) (Karkehabadi et al., 2008), TtGH61E (PDB entry 3EJA) (Harris et al., 2010), and TaGH61A (PDB entry 3ZUD) (Quinlan et al., 2011) was performed by using Superpose (Krissinel and Henrick, 2004) within the CCP4 package. Then this alignment was used as a profile in the program MAFFT (Katoh and Toh, 2008) for sequence alignment of all cellulose induced PMOs from *N. crassa* (Tian et al., 2009).

Multiple Sequence Alignments of PMO Subfamily Members

Homologs of PMO-2 and PMO-3 were found with a BLAST (Altschul et al., 1997) query of the PMO sequence against a database of predicted fungal proteins from finished and ongoing fungal genome projects. Representative basidiomycete and ascomycete sequences were chosen and each sequence was truncated at the N-terminus to remove the signal peptide so that the sequences started with His1. The C terminus was also truncated so that the last residue was an absolutely conserved proline (Pro218 in PMO-2 and Pro220 in PMO-3). Multiple sequence alignments were carried out using T-COFFEE software package with default parameters (Notredame et al., 2000). Each subfamily was defined by the group of PMOs that are orthologs for PMO-2, NCU02240, and PMO-3 in other fungi with a sequence identity higher than 50%.

Intact Mass Spectrometry of Native Proteins

Mass spectra were acquired using an Agilent 6510 quadrupole time-of-flight (Q-Tof) mass spectrometer. Crystals of *N. crassa* PMO-2 and PMO-3 were collected and washed in crystallization well solution and then dissolved in water. One microliter of a 1 μM protein sample was injected using an autosampler. Proteins were analyzed using an Agilent HPLC-chip cube and chip (G4240-62001) that contained Zorbax 300SB-C18 enrichment (4 mm; 40 nl) and analytical columns (34 × 75 μm). Solvent A was 97% water: 3% acetonitrile: 0.1% formic acid. Solvent B was 95% acetonitrile: 5% water: 0.1% formic acid. Protein was eluted using a gradient of 3% to 95% B over 9 min. Instrument parameters during data acquisition were as follows: HPLC-Chip voltage 1.85 kV, gas temperature 300°C, and drying gas flow 4 l/min. Samples were acquired at 1.03 spectra/s and 975.2 ms/spectrum with a range of 300–3,000 m/z. External mass calibration was performed immediately prior to measurements. Deconvolution of mass spectra was performed with Agilent MassHunter software.

Docking of PMOs on Crystalline Cellulose

A CIF file of cellulose Iβ containing the asymmetric unit coordinates (Nishiyama et al., 2002) was converted to PDB format by using Open Babel (O'Boyle et al., 2011). The cellulose crystalline matrix was generated with the SuperSym plugin in PyMol (http://www.pymol.org). For manual docking, the structures of CBM1, HjCel61B, TaGH61A, and *N. crassa* PMO-3 each were placed above the cellulose (100) face and oriented so that the protein planar surface was parallel to the (100) face and the stacking area between the aromatic rings and the pyranose units was maximized.

Modeling of Electron Transport Pathways in PMOs

Docking of the CDH heme domain from the white rot fungus *Phanerochaete chrysosporium* (PDB entry 1D7B) (Hallberg et al., 2000) onto *N. crassa* PMO-3 was performed with ZDOCK (Chen et al., 2003).

Sequence Alignment of CBM33s

The sequences of CBP21 from *Serratia marcescens* (Vaaje-Kolstad et al., 2005) and two other CBM33s (PDB entries 2XWX and 3UAM) were aligned based on their structures by using Superpose (Krissinel and Henrick, 2004) within the CCP4 package (Collaborative Computational Project, Number 4, 1994). Subsequently, CelS2 from *Streptomyces coelicolor* was added to the alignment by using the program MAFFT (Katoh and Toh, 2008).

Figure Preparation

All of the structural figures presented in this manuscript were prepared with the molecular visualization software PyMol (http://www.pymol.org).

ACCESSION NUMBERS

The structure models and structure factors for PMO-2 and PMO-3 have been deposited with the PDB as entries 4EIR and 4EIS, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and one table and can be found with this article online at doi:10.1016/j.str.2012.04.002.

ACKNOWLEDGMENTS

We thank G. Meigs and J. Holton for help with data measurement at beamline 8.3.1 at the Advanced Light Source, S. Bauer for technical assistance with LC-MS, and J. Klinman, C. Chang and V. Vu for helpful discussions regarding PMO reaction mechanisms. This work was funded by a grant from the Energy Biosciences Institute to J.H.D.C. and M.A.M. and the National Cancer Institute grant CA92584 for beamline 8.3.1 at the ALS.

Received: January 20, 2012
Revised: March 20, 2012
Accepted: April 4, 2012
Published online: May 10, 2012

REFERENCES

Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. (2002). PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D Biol. Crystallogr. *58*, 1948–1954.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.

Beeson, W.T., Phillips, C.M., Cate, J.H., and Marletta, M.A. (2012). Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. J. Am. Chem. Soc. *134*, 890–892.

Berka, R.M., Grigoriev, I.V., Otillar, R., Salamov, A., Grimwood, J., Reid, I., Ishmael, N., John, T., Darmond, C., Moisan, M.C., et al. (2011). Comparative genomic analysis of the thermophilic biomass-degrading fungi Myceliophthora thermophila and Thielavia terrestris. Nat. Biotechnol. *29*, 922–927.

Blanksby, S.J., and Ellison, G.B. (2003). Bond dissociation energies of organic molecules. Acc. Chem. Res. *36*, 255–263.

Boraston, A.B., Bolam, D.N., Gilbert, H.J., and Davies, G.J. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem. J. *382*, 769–781.

Burmeister, W.P. (2000). Structural changes in a cryo-cooled protein crystal owing to radiation damage. Acta Crystallogr. D Biol. Crystallogr. *56*, 328–341.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res. *37* (Database issue), D233–D238.

Chen, R., Li, L., and Weng, Z. (2003). ZDOCK: an initial-stage protein-docking algorithm. Proteins *52*, 80–87.

Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. Section D *50*, 760–763.

Eastwood, D.C., Floudas, D., Binder, M., Majcherczyk, A., Schneider, P., Aerts, A., Asiegbu, F.O., Baker, S.E., Barry, K., Bendiksby, M., et al. (2011). The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. Science *333*, 762–765.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Espagne, E., Lespinet, O., Malagnac, F., Da Silva, C., Jaillon, O., Porcel, B.M., Couloux, A., Aury, J.M., Ségurens, B., Poulain, J., et al. (2008). The genome sequence of the model ascomycete fungus Podospora anserina. Genome Biol. *9*, R77.

Fernandes, A.N., Thomas, L.H., Altaner, C.M., Callow, P., Forsyth, V.T., Apperley, D.C., Kennedy, C.J., and Jarvis, M.C. (2011). Nanostructure of cellulose microfibrils in spruce wood. Proc. Natl. Acad. Sci. USA *108*, E1195–E1203.

Foote, C.S., Valentine, J.S., Greenberg, A., and Liebman, J.F. (1995). Active Oxygen in Chemistry (New York: Chapman & Hall).

Forsberg, Z., Vaaje-Kolstad, G., Westereng, B., Bunæs, A.C., Stenstrøm, Y., MacKenzie, A., Sørlie, M., Horn, S.J., and Eijsink, V.G. (2011). Cleavage of cellulose by a CBM33 protein. Protein Sci. *20*, 1479–1483.

Galazka, J.M., and Cate, J.H.D. (2011). Improving the bioconversion of plant biomass to biofuels: A multidisciplinary approach. Energ. Environ. Sci. *4*, 3329–3333.

Gray, H.B., and Winkler, J.R. (1996). Electron transfer in proteins. Annu. Rev. Biochem. *65*, 537–561.

Hallberg, B.M., Bergfors, T., Bäckbro, K., Pettersson, G., Henriksson, G., and Divne, C. (2000). A new scaffold for binding haem in the cytochrome domain of the extracellular flavocytochrome cellobiose dehydrogenase. Structure *8*, 79–88.

Harris, P.V., Welner, D., McFarland, K.C., Re, E., Navarro Poulsen, J.C., Brown, K., Salbo, R., Ding, H., Vlasenko, E., Merino, S., et al. (2010). Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. Biochemistry *49*, 3305–3316.

Jahn, H.A., and Teller, E. (1937). Stability of polyatomic molecules in degenerate electronic states. I. Orbital degeneracy. Proc. R. Soc. Lon. Ser. A *161*, 220–235.

Karkehabadi, S., Hansson, H., Kim, S., Piens, K., Mitchinson, C., and Sandgren, M. (2008). The first structure of a glycoside hydrolase family 61 member, Cel61B from Hypocrea jecorina, at 1.6 A resolution. J. Mol. Biol. *383*, 144–154.

Katoh, K., and Toh, H. (2008). Recent developments in the MAFFT multiple sequence alignment program. Brief. Bioinform. *9*, 286–298.

Klinman, J.P. (2006). The copper-enzyme family of dopamine beta-monooxygenase and peptidylglycine alpha-hydroxylating monooxygenase: resolving the chemical pathway for substrate hydroxylation. J. Biol. Chem. *281*, 3013–3016.

Kraulis, J., Clore, G.M., Nilges, M., Jones, T.A., Pettersson, G., Knowles, J., and Gronenborn, A.M. (1989). Determination of the three-dimensional solution structure of the C-terminal domain of cellobiohydrolase I from Trichoderma reesei. A study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. Biochemistry *28*, 7241–7257.

Krissinel, E., and Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. D Biol. Crystallogr. *60*, 2256–2268.

Langston, J.A., Shaghasi, T., Abbate, E., Xu, F., Vlasenko, E., and Sweeney, M.D. (2011). Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. Appl. Environ. Microbiol. *77*, 7007–7015.

Larsson, P.T., Wickholm, K., and Iversen, T. (1997). A CP/MAS C-13 NMR investigation of molecular ordering in celluloses. Carbohydr. Res. *302*, 19–25.

Leahy, D.J., Aukhil, I., and Erickson, H.P. (1996). 2.0 A crystal structure of a four-domain segment of human fibronectin encompassing the RGD loop and synergy region. Cell *84*, 155–164.

Leiros, H.K.S., McSweeney, S.M., and Smalås, A.O. (2001). Atomic resolution structures of trypsin provide insight into structural radiation damage. Acta Crystallogr. D Biol. Crystallogr. *57*, 488–497.

Leiros, H.K.S., Timmins, J., Ravelli, R.B.G., and McSweeney, S.M. (2006). Is radiation damage dependent on the dose rate used during macromolecular crystallography data collection? Acta Crystallogr. D Biol. Crystallogr. *62*, 125–132.

Matthews, J.F., Skopec, C.E., Mason, P.E., Zuccato, P., Torget, R.W., Sugiyama, J., Himmel, M.E., and Brady, J.W. (2006). Computer simulation studies of microcrystalline cellulose Ibeta. Carbohydr. Res. *341*, 138–152.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.

Nishiyama, Y., Langan, P., and Chanzy, H. (2002). Crystal structure and hydrogen-bonding system in cellulose Ibeta from synchrotron X-ray and neutron fiber diffraction. J. Am. Chem. Soc. *124*, 9074–9082.

Notredame, C., Higgins, D.G., and Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. *302*, 205–217.

O'Boyle, N.M., Banck, M., James, C.A., Morley, C., Vandermeersch, T., and Hutchison, G.R. (2011). Open Babel: An open chemical toolbox. J. Cheminform. *3*, 33.

Ohm, R.A., de Jong, J.F., Lugones, L.G., Aerts, A., Kothe, E., Stajich, J.E., de Vries, R.P., Record, E., Levasseur, A., Baker, S.E., et al. (2010). Genome sequence of the model mushroom Schizophyllum commune. Nat. Biotechnol. *28*, 957–963.

Phillips, C.M., Beeson, W.T., Cate, J.H., and Marletta, M.A. (2011). Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by Neurospora crassa. ACS Chem. Biol. *6*, 1399–1406.

Quinlan, R.J., Sweeney, M.D., Lo Leggio, L., Otten, H., Poulsen, J.C., Johansen, K.S., Krogh, K.B., Jørgensen, C.I., Tovborg, M., Anthonsen, A., et al. (2011). Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. Proc. Natl. Acad. Sci. USA *108*, 15079–15084.

Saul, F.A., and Poljak, R.J. (1992). Crystal structure of human immunoglobulin fragment Fab New refined at 2.0 A resolution. Proteins *14*, 363–371.

Solomon, E.I., Ginsbach, J.W., Heppner, D.E., Kieber-Emmons, M.T., Kjaergaard, C.H., Smeets, P.J., Tian, L., and Woertink, J.S. (2011). Copper dioxygen (bio)inorganic chemistry. Faraday Discuss. *148*, 11–39, discussion, 97–108.

Taylor, C.B., Talib, M.F., McCabe, C., Bu, L., Adney, W.S., Himmel, M.E., Crowley, M.F., and Beckham, G.T. (2012). Computational investigation of glycosylation effects on a family 1 carbohydrate-binding module. J. Biol. Chem. *287*, 3147–3155.

Tian, C., Beeson, W.T., Iavarone, A.T., Sun, J., Marletta, M.A., Cate, J.H., and Glass, N.L. (2009). Systems analysis of plant cell wall degradation by the model filamentous fungus Neurospora crassa. Proc. Natl. Acad. Sci. USA *106*, 22157–22162.

Vaaje-Kolstad, G., Houston, D.R., Riemen, A.H., Eijsink, V.G., and van Aalten, D.M. (2005). Crystal structure and binding properties of the Serratia marcescens chitin-binding protein CBP21. J. Biol. Chem. *280*, 11313–11319.

Vaaje-Kolstad, G., Westereng, B., Horn, S.J., Liu, Z., Zhai, H., Sørlie, M., and Eijsink, V.G. (2010). An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. Science *330*, 219–222.

Wilson, D.B. (2009). Cellulases and biofuels. Curr. Opin. Biotechnol. *20*, 295–299.